Article details: 2014-0047	
Title	Treatment of overweight/obesity in children and youth: a systematic review with meta-analyses
Authors	Leslea Peirson, Donna Fitzpatrick-Lewis, Katherine Morrison, Rachel Warren, Muhammad Usman Ali, Parminder Raina
Reviewer 1	**Mandy Ho**
Institution	The Children's Hospital at Westmead Clinical School, University of Sydney, Sydney, New South Wales, Australia
General comments	Comments to authors It has been highlighted in the introduction that this review distinguished from others in the way that it will provide evidence for the effectiveness of child obesity interventions appropriate for primary care practitioners. However, this aim was not carried through well. The study questions have not been clearly defined. The rationales of the study and the clinical implications of the findings have not been adequately discussed. Specific comments 1. Page 3, last paragraph, authors need to make clear to the readers and state in the abstract as well whether this review is an update of the US Preventive Services Task Force's review. Did the current review use the same search strategies as the previously mentioned review? Sargent and colleagues have published a systematic review on primary care interventions for child obesity which I believe it may be more relevant to the scope of the current review (Sargent, G. M., L. S. Pilotto, and L. A. Baur. "Components of primary care interventions to treat childhood overweight and obesity: a systematic review of effect." Obesity reviews 12.5 (2011): e219-e235.). 2. Figure 1, the flow diagram could be improved, please refer to the PRISMA guideline. It is not clear why there's a "child obesity treatment update search: 2041". What does the box "included in US Preventive Services Task Force Review: 15" mean? Were these 15 studies not picked up by the initial search? 3. Box 1, Population: it's not clear what you mean by "met previously accepted criteria for overweight based on ideal body weight'? 4. Box 1, Settings: how do you define "generalizable to Canadian primary care or feasible for conducting in or referral from primary care" 5. Box 2, inclusion criteria #5 "no restrictions on study design, comparison group, weight outcome reporting, or timing of assessment were applied to studies that reported data for harms of treatment" was confusing. 6. Box 2, exclusion criteria #1, what does the "built environment" refer to? Does it include the "home environment" and/or "school environment"? Need to justify why only included studies of orlistat but not other drugs. 7. Page 4, Study Selection, Please state explicitly that the "Grading of Recommendations Assessment, Development and Evaluation System" is adapted from Cochrane Review. 8. Page 4-5, Data Analysis, what meta-analysis program was used for synthesizing the data? 9. Page 5, line 18, it's not clear what do you mean by "to help interpret SMD, values were converted to BMI and BMI z-score units". Line 27-32, need to justify why using the standard deviation value of one particular study to reflect the among-person variation? If there're enough studies, suggest use mean different instead of standardized mean difference. Has publication bias been assessed? 10. Results, page 6, Line 23 to 33, I believe these information should be included in the method instead. Line 33, again, why there's an update search which added 2041 citations? 11. Results, characteristics and quality of the included studies have not been adequately discussed. Also, need to comment on the heterogeneity of the results. 12. Table 2, the table is very complex and dense. Need to describe in the text (methods section) how was the "quality of evidence rating" assessed. 13. Table 3, I believe forest plots would be a more appropriate way for presenting the results. 14. Page 7, Line 5-13, I found this statement very confusing, same for line 42 to 46. 15. Page 9, line 3-4, It's interesting that only "laps or stages of the multi-stage fitness test" was mentioned here. Did the included studies evaluate physical fitness using other measures? This has not been mentioned in the methods. 16. Page 9, second paragraph, how was quality of life assessed? Table 3, what did the quality of life scores mean? What're the ranges of the quality of life scores? Does a higher score reflect a better quality of life or the opposite? 17. Page 9, Harms, I found this section very confusing. 18. Page 9, Interpretation, Line 44-53, references were required for this statement. 19. Page 11, 1st paragraph, need to be further elaborated. Could the difference be related to different intervention intensity or the different background/level of obesity of the study participants? 20. What are the implications for primary care practitioners?
Reviewer 2	**Iris Mabry-Hernandez, MD, MPH**
Institution	Agency for Healthcare Research and Quality, Rockville, MD

General comments	General comments: Systematic review/meta-analysis to look at if primary care relevant treatment interventions lead to short-term or sustained weight stabilization, reduction, or other health benefits. Introduction: There is a lack of a general discussion of what types interventions may work with children/adolescents and their families. No discussion of orlistat and its indications of use. (p.3 lines 6-39) Although the authors state why this review is unique, they don't explicitly state why it may be important for clinician to have effective interventions conducted in the office or offered in parallel with "community resources." Please define "community resources." Why aren't faith-based organizations included? The authors mention that several systematic reviews have been published but they don't mention the results of these other reviews. It seems as though there should be more than 1 key question- at least one about benefits and another about harms. (p.3 lines 6-39) Methods: With regards to inter-rater conflicts, who was involved in the discussions? (p.4, line 47) I would like to see more of a discussion about heterogeneity, including levels of heterogeneity. Why no funnel plots (too much heterogeneity)? (p.5 line49) There is no detailed discussion of the types of interventions examined, components of the various interventions, how intensity of interventions may play a role, how "primary care relevant " interventions were defined. QOL should be better defined in text. How was it measured? Validation should be mentioned. (p.9 line 8)? Figure 1 should include reasons for exclusions at all levels. (p.20) Results: A more complete table of the study characteristics should be included in the appendix. (p.22) Please mention what the gender breakdown is (e.g., 50% female, baseline BMI, percent overweight vs. obese, a few more details on intervention type, settings). I am curious as to what type of individually –based intervention can address the needs of 2 year olds in it? (p.22, line 12 is one example) Figure 2-consider having forest plots with some of the subgroups. Interpretation: I didn't see a discussion of generalizability of the results? How diverse were the participants? There is only 1 sentence for future research needs (p.11, line 51)- please expand. Iris Mabry-Hernandez, MD, MPH Agency for Healthcare Research and Quality Rockville, MD I not have any conflicts of interest to disclose.
Author response	Reviewer #1 (MH) Comments to Authors' Revisions/Response General Comments It has been highlighted in the introduction that this review distinguished from others in the way that it will provide evidence for the effectiveness of child obesity interventions appropriate for primary care practitioners. However, this aim was not carried through well. The study questions have not been clearly defined. The rationales of the study and the clinical implications of the findings have not been adequately discussed. AR#26 This review followed a peer reviewed protocol. We believe some of the reviewer's concerns have been addressed through the revisions made to the manuscript. Specific Comments 1. Page 3, last paragraph, authors need to make clear to the readers and state in the abstract as well whether this review is an update of the US Preventive Services Task Force's review. Did the current review use the same search strategies as the previously mentioned review? Sargent and colleagues have published a systematic review on primary care interventions for child obesity which I believe it may be more relevant to the scope of the current review (Sargent, G. M., L. S. Pilotto, and L. A. Baur. "Components of primary care interventions to treat childhood overweight and obesity: a systematic review of effect." Obesity reviews 12.5 (2011): e219-e235.).

	AR#27 We considered the studies included in the USPSTF review against our inclusion criteria. We did not bring forward all of the studies they included. Our search strategy was based on the USPSTF strategy (see Response #1 for details). 2. Figure 1, the flow diagram could be improved, please refer to the PRISMA guideline. It is not clear why there's a "child obesity treatment update search: 2041". What does the box "included in US Preventive Services Task Force Review: 15" mean? Were these 15 studies not picked up by the initial search? AR#28 Flow diagram was revised as per Editor's feedback (see response #24). We believe our flow diagram does include the PRISMA elements. In the methods we clarify that in addition to using the USPSTF's search strategy to update the evidence we planned to bring forward any studies from the USPSTF 2010 review that met our inclusion criteria. As noted in the search results section, to assess these 15 studies for relevance they were considered in the selection process (and inserted in the flow diagram) at the point of full text screening. Since our search was an update (dates subsequent to the USPSTF parameters), these 15 studies would not have been picked up. 3. Box 1, Population: it's not clear what you mean by "met previously accepted criteria for overweight based on ideal body weight'? AR#29 This phrase was deleted. 4. Box 1, Settings: how do you define "generalizable to Canadian primary care or feasible for conducting in or referral from primary care" AR#30 The explanation provided in the Setting section of Box 1 has been expanded. 5. Box 2, inclusion criteria #5 "no restrictions on study design, comparison group, weight outcome reporting, or timing of assessment were applied to studies that reported data for harms of treatment" was confusing. AR#31 This point was rephrased to improve clarity. 6. Box 2, exclusion criteria #1, what does the "built environment" refer to? Does it include the "home environment" and/or "school environment"? Need to justify why only included studies of orlistat but not other drugs. AR#32 Added examples of changes to the built environment. Added point that orlistat is only drug approved for weight loss by Health Canada. 7. Page 4, Study Selection, Please state explicitly that the "Grading of Recommendations Assessment, Development and Evaluation System" is adapted from Cochrane Review. AR#33 We cannot make the recommended statement. GRADE was developed by the GRADE Working Group which is not part of Cochrane, although there are strong connections between these entities. Cochrane has adopted the GRADE approach for its reviews. 8. Page 4-5, Data Analysis, what meta-analysis program was used for synthesizing the data? AR#34 A statement was added to the Data Analysis section indicating that Review Manager, STATA and GRADEpro were the programs used for analyses. 9. Page 5, line 18, it's not clear what do you mean by "to help interpret SMD, values were converted to BMI and BMI z-score units". Line 27-32, need to justify why using the standard deviation value of one particular study to reflect the among-person variation? If there're enough studies, suggest use mean different instead of standardized mean difference. Has publication bias been assessed? AR#35 We agree this additional information could be confusing to readers. These conversions were originally performed to support the guideline group using the evidence review. We removed these conversions from the manuscript. The information on lines 27-32 was no longer relevant and was removed from the manuscript. In addition to the SMD for BMI/BMI-Z calculated for 30 studies we performed a meta-analysis on the 21 studies that reported BMI; for this analysis we were able to use MD. Yes, when the number of studies was ≥10 (as per Cochrane Handbook) reporting bias was examined as a component of the GRADE assessment – we used STATA to perform the Egger's Test. 10. Results, page 6, Line 23 to 33, I believe these information should be included in the method

	instead. Line 33, again, why there's an update search which added 2041 citations? AR#36 See response #24. 11. Results, characteristics and quality of the included studies have not been adequately discussed. Also, need to comment on the heterogeneity of the results. AU#37 See response #17. Also, the summary of risk of bias ratings for each study was moved from the supplemental e-File to the main manuscript (appears as Table 2). Regarding heterogeneity see response #4. 12. Table 2, the table is very complex and dense. Need to describe in the text (methods section) how was the "quality of evidence rating" assessed. AR#38 We removed the details for conversions to BMI and BMI-z (see response #35) which removed some of the complexity/density of the table. Details were added to the methods section and in a footnote in each data table elaborating on GRADE assessment criteria and ratings. 13. Table 3, I believe forest plots would be a more appropriate way for presenting the results. AR#39 We believe it is appropriate and efficient to provide the results for multiple outcomes in a single table as opposed to several forest plots. The table contains and organizes the information available in the forest plots plus provides the GRADE rating for each sub-body of evidence. 14. Page 7, Line 5-13, I found this statement very confusing, same for line 42 to 46. AR#40 These statements were re-worded. 15. Page 9, line 3-4, It's interesting that only "laps or stages of the multi-stage fitness test" was mentioned here. Did the included studies evaluate physical fitness using other measures? This has not been mentioned in the methods. AR#41 We selected the validated multi-stage fitness test (Leger's test) as our measure for physical fitness. In the Methods section we added a note in Box 1 (outcomes) to clarify. 16. Page 9, second paragraph, how was quality of life assessed? Table 3, what did the quality of life scores mean? What're the ranges of the quality of life scores? Does a higher score reflect a better quality of life or the opposite? AR#42 We added information about the QOL tools used in the studies and their validation. 17. Page 9, Harms, I found this section very confusing. AR#43 This section was edited to improve clarity. 18. Page 9, Interpretation, Line 44-53, references were required for this statement. AR#44 References were added. 19. Page 11, 1st paragraph, need to be further elaborated. Could the difference be related to different intervention intensity or the different background/level of obesity of the study participants? AR#45 We added a few points to elaborate on this observation. 20. What are the implications for primary care practitioners? AR#46 A statement was added to address this comment. Reviewer #2 (IM-H) Comments to Authors Authors' Revisions/Response # General Comments Systematic review/meta-analysis to look at if primary care relevant treatment interventions lead to short-term or sustained weight stabilization, reduction, or other health benefits. AR#47 No action required Specific Questions

	Introduction 1. There is a lack of a general discussion of what types interventions may work with children/adolescents and their families. No discussion of orlistat and its indications of use. (p.3 lines 6-39) AR#48 Points were added to the introduction section as requested. Introduction 2. Although the authors state why this review is unique, they don't explicitly state why it may be important for clinician to have effective interventions conducted in the office or offered in parallel with "community resources." Please define "community resources." Why aren't faith-based organizations included? The authors mention that several systematic reviews have been published but they don't mention the results of these other reviews. It seems as though there should be more than 1 key question- at least one about benefits and another about harms. (p.3 lines 6-39) AR#49 We rephrased some of the interpretation and acknowledge in our conclusion that low intensity behavioural interventions may be feasible for implementation by family physicians/clinicians. The term "community resources" has been removed. The settings for intervention are clarified in Box 1. Faith-based programs (i.e., religious or spiritual content) were excluded as being beyond the scope of this review that focuses on primary care relevant interventions. We did not exclude programs that were delivered in faith-based settings (e.g., using a church basement for meetings). Results of other reviews are covered later in the interpretation section. Details were added that clarify the intent of the review to examine benefits and harms of interventions and features of efficacious interventions. Methods 3. With regards to inter-rater conflicts, who was involved in the discussions? (p.4, line 47) AR#50 Details added to clarify that conflicts were resolved through discussion between raters and if necessary through consultation with other members of the review team. Methods 4. I would like to see more of a discussion about heterogeneity, including levels of heterogeneity. Why no funnel plots (too much heterogeneity)? (p.5 line49) AR#51 See response #4. We assessed publication bias using the Egger's test as part of our GRADE rating of overall quality of evidence. Methods 5. There is no detailed discussion of the types of interventions examined, components of the various interventions, how intensity of interventions may play a role, how "primary care relevant" interventions were defined. AR#52 See response #17. The last point about defining primary care relevant is clarified in Box 1. Methods 6. QOL should be better defined in text. How was it measured? Validation should be mentioned. (p.9 line 8)? AR#53 We added information about the QOL tools used in the studies and their validation. Methods 7. Figure 1 should include reasons for exclusions at all levels. (p.20) AR#54 PRISMA requests reasons are provided for exclusions at full text level only. We provide this information in our flow diagram. Results: 8. A more complete table of the study characteristics should be included in the appendix. (p.22) Please mention what the gender breakdown is (e.g., 50% female, baseline BMI, percent overweight vs. obese, a few more details on intervention type, settings). I am curious as to what type of individually –based intervention can address the needs of 2 year olds in it? (p.22, line 12 is one example) AR#55 A table was added to the supplemental e-File that contains characteristics of the individual studies. The mean age at baseline ranged from 5 to 16 – there were no treatment programs aimed at very young children. For each study in the main manuscript Table 1 (Summary of Characteristics of Included Studies) we replaced the age range designation (2-12 or 13-18) with the mean baseline

	age. The age ranges 2 to 12 and 13 to 18 were used by the USPSTF for analyses. To be consistent we categorize studies according to these general age groups for our sub-group analysis (Table 3). Results: 9. Figure 2-consider having forest plots with some of the subgroups. AR#56 See Response #39. If the editors want to include more forest plots in the manuscript we can provide them. Interpretation: 10. I didn't see a discussion of generalizability of the results? How diverse were the participants? There is only 1 sentence for future research needs (p.11, line 51)- please expand. AR#57 The generalizability of the results are represented by the overall GRADE quality of evidence (High, Moderate, Low or Very Low) and the diversity of the population is assessed using the directness domain (PICO criteria) of the GRADE assessment. The section on future research was expanded.